THE UNIVERSITY OF ALBERTA

THE MAC AS A CULTURE-REDUCED MEASURE OF INTELLECTUAL POTENTIAL

by

JAMES E. CARLSON

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

APRIL, 1966

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned hereby certify that they have read and recommended to the Faculty of Graduate Studies for acceptance, a thesis entitled, "The MAC as a Culture-Reduced Measure of Intellectual Potential", submitted by James E. Carlson in partial fulfillment of the requirements for the degree of Master of Education.

ABSTRACT

This study was designed to evalute the MAC as an instrument for the measurement of intellectual potential with a minimum of cultural bias.

The MAC was administered to samples of Eskimos from Inuvik and Tuktoyaktuk, Northwest Territories and Indians and Metis from Faust, Alberta as a part of large batteries of tests of intellectual ability. In addition, the MAC and two conventional group intelligence tests were administered to a sample of White urban children.

The main criteria selected for culture-reduced tests, or tests designed to measure intellectual potential with minimal cultural bias were: such a test should have high loadings on a general intellectual ability factor and low loadings on group factors as determined by factor analysis; and such a test should be less biased culturally than are conventional group intelligence tests.

Principal axes factor analysis showed that the MAC adequately met the criteria of high general factor loadings and low group factor loadings for three out of four samples. Tests of the significance of differences between means showed that the MAC was less culturally biased than were two conventional intelligence tests for age 12½ to 15½ Eskimos and Indian-Metis but not for Eskimos and Indian-Metis of ages 9 to 12.

Some estimates of the reliability of the MAC were calculated and it was suggested that the test be expanded by the addition of more items and that the format be changed slightly.

ACKNOWLEDGEMENTS

The writer wishes to express his appreciation to the following persons for their assistance during the development of this thesis.

Recognition is due, in particular, to Dr. R. S. MacArthur, without whose constant guidance and encouragement this work would not have been possible, and to the other members of the committee, Drs. D. Fair and C. W. Hobart for their helpful suggestions and constructive criticism. For their kind permission to use adaptations of their tests, thanks are due to Personnel Research of the South African Council for Scientific and Industrial Research, and to the Deputy Director of the Institute, Mr. A. O. H. Roberts. Special thanks are due to Dr. MacArthur for the use of data collected for his Mental Ability Surveys in Faust, Alberta, and Inuvik and Tuktoyaktuk, Northwest Territories.

The writer is also indebted to the principals, teachers and school board officials of the schools involved, since their cooperation made this study possible, and to Mr. Rod Conklin and Mr. K. O'Bryan for their helpful suggestions while the thesis was being written.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

THE THESIS PROBLEM AND ITS BACKGROUND

Investigators dealing with the influence of social class and cultural factors on measured intelligence have generally found that the mean Intelligence Quotient of middle-class, white North Americans from urban centers is significantly higher than that of persons from non-white races or lower social classes. However, there has been considerable disagreement among investigators about the reasons why cultural differences are found when measuring mental ability. There is even less agreement with respect to what, if anything, should be done about these differences.

Most present-day researchers appear to attribute some of the differences found when measuring intellectual ability in people from different cultures to a cultural bias inherent in the instruments being used. It is generally recognized that contemporary group tests designed to measure intellectual ability are biased against persons from cultures other than that of middle-class, white, urban European or North American cultures. However, there are differences of opinion with respect to the significance of this bias when dealing with persons from such cultures.

The argument is sometimes heard that, since the purpose of measuring the intellectual ability of persons from "culturally-deprived" backgrounds is to arrive at an estimate of their possible

success in a particular type of society, that ability should be measured within the context of the culture of that society. However, this argument assumes that the amount of bias will be constant for all individuals measured, a possibility that is highly unlikely.

When a test is used outside of the culture for which it was designed, there is an implicit assumption that, although the instrument being used may not accurately assess a person's level of ability, discrimination between persons of high and low levels of ability will be accurate. Since predictive efficiency is a function of the reliability of the instrument being used, and conventional group intelligence tests are admittedly unreliable when used outside of the culture for which they were designed, the validity of these instruments for such use is highly questionable.

Immigrants, when they first enter the North American society, usually undergo no assessment of abilitites. Instead, they are allowed to find work at unskilled jobs until they become assimilated into the adopted culture with the hope that those with the most ability will find their way into higher level occupations. Studies of immigrants arriving in the United States have indicated that one generation may pass before this form of upward mobility takes place.

The ideal solution to the problem of measuring the intellectual ability of persons from a variety of cultural settings would be to develop a test that is completely "culture-free" - a rather remote possibility. However, the development of a test that is less culturally biased than is the conventional group intelligence test may be within the

realm of possibility. Such a test should employ stimulus symbols that are relatively familiar in a variety of cultures and should measure a person's "present capacity for future development of intelligent behavior" (West & MacArthur, 1964, p. 18) as opposed to his present level of proficiency. A test of this nature could conceivably sample cognitive behavior with a reduction in bias due to highly verbal instructions, questions, and responses, and unfamiliar symbols. If a test of intellectual potential can be developed that is less biased than the conventional test, then the person entering a culture that is new to him can be placed in an occupation more closely commensurate with his potential ability than would be the case without such a test. At the same time, a program of adaptive treatment could be instituted that would be geared to the individual's measured potential and which had the goal of bringing his present level of functioning closer to this potential ability.

The characteristics of instruments that can be used for the purpose of measuring intellectual potential with a minimum of cultural bias has been the subject of research by MacArthur and his associates at the University of Alberta for several years. Promising instruments have been identified, studied, modified and evaluated against criteria for measuring intellectual potential with minimal cultural bias. Recently, as part of this ongoing research, MacArthur has experimented with the MAC, an instrument adapted from tests used in preliminary form by the National Institute for Personnel Research

in South Africa. The MAC appears to be a promising instrument for the above-mentioned purpose in that it meets criteria such as simplicity of symbols used as stimuli, low level of language required for understanding, and administrative practicality. However, there are several other criteria against which the MAC must be evaluated before it can be used.

The general purpose of this study is to investigate the MAC as an instrument for the assessment of intellectual potential with a minimum of cultural bias and to evaluate it in a statistical and psychological sense against selected criteria for such instruments. More specifically, the purpose is to evalute the MAC against the selected criteria for samples of Eskimo children from Inuvik and Tuktoyaktuk, Northwest Territories, and Indian and Metis children from Faust, Alberta. Comparisons will be made between these samples and a sample of white children from an Edmonton school. Each sample will be divided into two age groups: nine to twelve years, and twelve and one-half years to fifteen and one-half years.

CHAPTER II

REVIEW OF RELATED LITERATURE

I. Measurement of Intelligence

Vernon (1965b) points out that most contemporary psychologists appear to have arrived at the conclusion that the "intelligence" that we measure is the resultant of interaction between inherited potential and environmental conditions. Similarly, West and MacArthur postulate that "individual differences in measured intelligence result from the interaction of heredity and environment in a multiplicative manner" (West & MacArthur, 1964, p. 17). They also point out that the word "intelligence" is often used in two related ways: innate substratum of predispositions (Hebb's Intelligence A), and present level of proficiency (Hebb's Intelligence B). A third usage is proposed which West and MacArthur call "Intelligence A'" or "Present capacity for future development" (1964, p. 17).

When measuring intellectual potential, MacArthur (1964), postulates that it is Intelligence A' that is being measured, and that the most parsimonious estimate of this potential is a general intellectual ability factor, similar to Spearman's g but not solely innate. He reasons that environmental circumstances may not always allow the individual's mental abilities to develop to their full potential. However, through adaptive procedures, it may be possible

to increase the individual's present level of functioning so that it more closely approximates the level that might have been reached under more favorable environmental conditions. If this is so, then it may be possible to measure the individual's intellectual potential or capacity for future development.

Both MacArthur (1962a, 1962b, 1964) and Vernon (1956, 1965a, 1965b) view intelligence as being hierarchical in nature, with a general intellectual ability factor at the top of the hierarchy. Further down the hierarchy lie group factors of a more specific nature which can be further divided into more specific abilities. When MacArthur postulates that the general ability factor is the most parsimonious estimate of intellectual potential, an underlying premise is that the more specific abilities are more susceptible to differential development in differing cultures. The fact that child-rearing practices vary considerably from culture to culture is one of the reasons for this differential development. These varying practices are, in turn, at least partially caused by variations in technologies.

Thus, MacArthur and Vernon favor the model of intelligence proposed by British factor analysts. In contrast, many American psychologists such as Thurstone prefer to think of intelligence in terms of many specific abilities with no common factor running through all these abilities. It has been argued that the two positions are not as different as they may first appear to be. By starting with an

American type of factor analysis, with oblique factors, and then proceeding to take out second or third order factors (or higher order if necessary) it may be possible to arrive at the general ability factor which is similar to that which the British methods take out prior to specific factors.

For the purposes of this study, intelligence will be considered to be hierarchical in nature, as postulated by MacArthur and Vernon. It follows from the hierarchical model that, for cross-cultural measurement of intellectual potential, the best assessment can be made by sampling behavior of a cognitive nature through the use of a variety of stimuli that are likely to be familiar in a variety of cultures.

As cultural bias in traditional measures of intellectual ability has become more widely recognized, many attempts have been made to construct "culture-free" and "culture-fair" tests of intelligence. It is unfortunate that these terms have come into the literature, because they seem to imply completely unbiased tests. Most psychometricians, while recognizing cultural bias as a factor in testing, doubt that a completely unbiased test can be constructed. To refer to tests as being "culture-reduced" would be more meaningful than referring to them as "culture-free" or "culture-fair" instruments. The phrase "culture-reduced" in this context means less culturally biased than other tests.

II. Criteria for Instruments

In trying to construct tests that will measure intellectual ability while minimizing cultural bias, investigators have outlined criteria for such tests. Most of these outlines have some criteria in common but some are more comprehensive than others.

Hess (1955) followed two main criteria for choosing items for his instruments. The first criterion was that the problems should resemble real life situations rather than academic tasks. His second criterion was rather general, stating that no cultural advantage should be given any group by the problems, vocabulary of instructions, materials or motivation used.

Knapp (1960) studied the effects of time limits on performance on the Institute for Personality and Ability Testing, test of g, for one hundred adult male Mexicans and an equal number of adult male Americans. He used two forms of the test, administering one under power conditions and the other under speeded conditions, breaking the samples into enough groups in order to control for the order of administration of the two forms, as well as for the order of administration under power and speeded conditions. Knapp found that both groups tended to score higher under power conditions than under speeded conditions but that the difference was significantly greater for the Mexicans than for the Americans. This study would tend to support the hypothesis, put forward by many other researchers, that tests used with different cultural groups should be primarily power rather than speeded tests.

Yates' (1963) study also supports the use of power tests. Using two groups of university students and the Progressive Matrices test, he found that for about 13 per cent of the sample, the score after 40 minutes markedly underestimated the level that the subject could reach when given a longer time (60 minutes). Conceivably, the percentage would be higher for younger subjects since there would be fewer slow-working individuals in a university than in a school population.

Campbell (1963) argues that although there seems to be general agreement that items are "culture-free" or "culture-fair" if all subjects are acquainted with the item type and its general method of solution, familiarity with specific instances is also necessary if the items are not to be culturally biased.

Schwarz (1963) states that each examinee must know precisely what it is that he is required to do and that, in order to accomplish this degree of familiarity with the type of test item, he argues that overlearning must be allowed to take place through the inclusion of a large number of practice items. Another criterion advocated by Schwarz is that symbols be used whose denotations need not be understood in order to solve the problems.

Because cultural anthropologists have demonstrated that the topological equivalents of perceptual variables appear in designs made by the people of most cultures, Kidd (1965) postulated that these spatial forms might yield the most valid cross-cultural

measurement tasks. He argues that "culture-fair" items should therefore be based on elementary spatial concepts.

MacArthur (1964) has outlined an elaborate set of criteria for evaluation of instruments designed to measure intellectual ability in a variety of cultural settings. He has suggested the following nine criteria:

1. It should largely sample the broad factor of general intellectual ability running through a wide variety of European-American kinds of intellectual tasks, including arithmetic, linguistic, and reading achievement.
2. It should have negligible loadings on verbal, numerical, and other group factors. That is, it should sample general intellectual ability through simple basic symbols not very dependent upon particular previous learnings.
3. It should show less difference between cultures in cross-cultural administration than do alternative measures of intelligence.
4. It should show moderate relationship with current school achievement.
5. It should show evidence of stability under changed environmental conditions relative to that of alternative measures of intelligence.
6. It should minimize effects of test sophistication, providing plenty of appropriate practice experience, having directions depending little if at all upon language, and being unspeeded.
7. It should be reliable.
8. It should be practical and usable from an administrative standpoint.
9. When and if possible, it should show long-term validity as a predictor of success in intellectual tasks when appropriate <u>adaptive</u> intervening treatments have been employed.

(MacArthur, 1962a, pp. 5-6)

Some of the tests which MacArthur and his associates have found to meet these criteria less badly than do the conventional group

intelligence tests are: Raven's Progressive Matrices, the Safran Culturally Reduced Intelligence Test (SCRIT), the Institute for Personality and Ability Testing Test of g (Cattell), and certain subtests of the California Test of Mental Maturity (CTMM) and Lorge-Thorndike Intelligence Tests - Nonverbal. In addition to these tests MacArthur has been investigating for possible use the MAC, an adaptation of tests used in preliminary form by the National Institute for Personnel Research in South Africa.

MacArthur suggests the following as possible reasons that some tests appear to meet the criteria less badly than do conventional group intelligence tests:

> (a) they form something of an age-scale sampling stages in the development of human cognition, starting with perception-dominated items, and proceeding through reversible concrete operations, to propositional or formal operations; (b) they use as stimuli symbols, which while dependent on learning, are likely to be learned in a variety of cultures; (c) they are crudely programmed samples of learning-on-the-spot.
>
> (MacArthur, 1964, p. 6)

III. Related Research

Eells et. al. (1951) studied the intelligence test results of 5000 pupils in two age groups, 9 through 10 years and 13 through 14 years. The battery for the younger group consisted of the Otis Alpha, Henmon-Nelson, and Kuhlman-Anderson tests, and that for the older group of the Terman-McNemar, Otis Beta, California Short Form Test of Mental Maturity and Chicago Tests of Primary Mental Abilities. As a measure of social status they used the Index of Status

Characteristics. Their findings were that about 50 per cent of the items from the 9 and 10 year old battery and about 85 per cent of those from the 13 and 14 year old battery showed differences between high- and low-status groups large enough to be significant at the .01 level. More than one-third of the items from the tests given the younger pupils and about one-tenth of those from the tests given to the older pupils had status differences too small to be significant at the .05 level. Mean status differences were largest for verbal and smallest for picture, geometric design and stylized drawing items. A substantial number of the items having large status differences involved an "academic" type of vocabulary.

The statistical procedures used in this study have, however, been criticized in some quarters. Tyler (1963), for example, states that before they concluded that differences in measured intellectual ability between members of different social classes were the result of test bias, the authors of this study should have taken into consideration the reliabilities of the various tests, the equality of IQ units within a test and between tests, and the difficulty levels of the various tests.

Marquart and Bailey (1955) in a study involving 71 children found a high correlation between IQ scores on the Cattell and the Stanford Binet, and suggested that the Cattell does not measure anything different than that measured by the Binet. In addition, it was found that for children in the 4 to 7 year old age range, scores on the

Binet were less influenced by social class than were scores on Scale 1 of the Cattell. However, for children in the 8 to 15 year old age group, scores on Scale 2 of the Cattell were found to be less influenced by social class than were the Binet scores.

Hess (1955) attempted to develop a test of mental ability which would control cultural variables related to socio-economic status. He used the previously cited criteria in order to gather 24 test items, 8 of which were discarded because of inappropriate difficulty levels, scoring problems, or other administrative faults. He then drew a sample of 545 elementary school pupils between the ages of 6½ and 9½ years and divided them into four age groups (6½ years, 7½ years, 8½ years, and 9½ years) and three status groups (high-status white, low-status white, and low-status Negro) on the basis of fathers' occupations. Mental ages derived from standard intelligence tests were taken from school records. The mean mental ages of the high-status white groups were significantly higher than those of low-status groups at each age level. Similar status differences were found on achievement test scores. On the experimental test, differences between the high-status white and low-status Negro groups were significant at all age levels but differences between the two white groups were not significant. Hess concluded that socio-economic differences between high- and low-status samples in the United States are exaggerated by standard intelligence tests.

Knief & Stroud (1959) dealing with a sample of 344 fourth grade pupils in a Midwestern United States city calculated the correlation

coefficients between several tests and the Warner Index of Status Characteristics. The correlations found between social class scores and the various test scores were: Davis-Eells Games .309; Lorge-Thorndike - Verbal .304; Lorge-Thorndike - Nonverbal .323; Progressive Matrices .179.

Elley (1961) studied a total of nine "culture-reduced" intelligence tests and found that the Progressive Matrices best met all the criteria adopted (similar to MacArthur's nine). He also found that the Cattell and the Figure Analogies subtest (subtest 3) of the Lorge-Thorndike - Nonverbal fulfilled nearly all of the criteria. His sample consisted of 432 grade seven children in Edmonton and were shown to be representative of all grade 7 children in that city.

Anderson (1962) in a study of the relation of Lorge-Thorndike Intelligence Test scores to socio-economic status found that the scores on that test are highly related to the social class of the subject's family. However, he found that the mean verbal IQ for the lowest social class was significantly higher than the mean nonverbal IQ at the .05 level. This test (Nonverbal) has, however, been shown by other research to be less biased than conventional verbal intelligence tests (West, 1962). Anderson's sample consisted of 319 fifth graders and 279 seventh graders in the Greater Syracuse, New York area. The group sizes were 48, 213, and 337 from highest to lowest social classes respectively.

West (1962) using MacArthur's nine criteria, found that the following tests gave the most adequate assessments of intellectual potential with minimum cultural bias at the grade levels indicated: at the grade one level, the SCRIT, Lorge-Thorndike subtests 2 and 3, and Progressive Matrices; at the grade two and three levels, the CTMM Non-language, Lorge-Thorndike subtests 2 and 3, SCRIT and Progressive Matrices; at the grade five and six levels, the Progressive Matrices, Cattell and SCRIT; and at the grade seven and eight levels, the Lorge-Thorndike subtests 1 and 2 or total, Progressive Matrices, CTMM Non-language and Cattell. West's sample consisted of 126 Metis children in grades one to eight at Faust, Alberta and 135 Indians and Metis in grades one to eight at Fort Simpson, Northwest Territories.

West and MacArthur (1964) in a factor analytic investigation into the construct validity of nine culture reduced tests found that the Progressive Matrices, the Cattell and the SCRIT best met the criteria adopted, which were essentially the first four of MacArthur's nine. The sample consisted of 147 Metis children of both sexes who were attending school at Faust, Alberta.

In summary, research reported here tends to show that certain group measures of intellectual ability are less biased culturally than are the conventional, highly verbal, intelligence tests.

CHAPTER III

THE MAC TEST

As previously mentioned, the MAC has been adapted by MacArthur from tests used in preliminary form by the National Institute for Personnel Research in South Africa. MacArthur felt that the symbols used as stimuli on many of the so-called "culture-free" tests were not likely to be equally familiar in a variety of cultures. Therefore the development of a test which used symbols of a more common and simple nature seemed desirable.

The MAC test as used in the present study is in a preliminary stage of development for use only in exploratory studies. Previous to this study there existed no data concerning the validity, reliability, difficulty level or cultural bias of the test. One of the purposes of this investigation is to determine whether or not further exploration of the MAC is warranted.

In its present form the MAC consists of three parts, the MAC1, MAC2 and MAC3. Because the MAC3 was deemed too difficult for the majority of the subjects in this study, only the MAC1 and MAC2 were used. Therefore the term MAC, as used in this study, shall refer to the MAC1 and MAC2.

In an attempt to reduce lack of familiarity with the item types among subjects, items one to five of the MAC1 were carefully taught as the first step in the administration of the MAC. The

items were taught one by one, and individual help was given to those subjects who had difficulty finding or recording correct responses. The responses of all testees were checked individually. Next, subjects were instructed to complete items six through ten of the MAC1. These items were corrected before the testing portion of the administration began and individual help was again given. The remaining ten items of the MAC1 were then used as test items. Upon completion of the MAC1 items one and two of the MAC2 were used as additional teaching items, and, finally, the remaining eighteen items were given as test items. These eighteen items were used as the MAC2 for the analyses of this study. A copy of the instructions used in the administration of the MAC can be seen in Appendix C.

The items of the MAC1 and MAC2 are placed in two separate booklets with five items to a page in each booklet. The answer sheets are inserted in the booklets in such a way that, as each page is turned, the spaces designated for recording the answers selected for the items of the following page are beside the appropriate items. Each item of the MAC is a series of triangles, circles, squares, or some combination of these symbols. Three different sizes of symbol and three colours are used. In order to prevent confusion to colour blind subjects, different types of shading are used with each colour. The task given to the testee is to choose the next two figures in the series from a chart of all the possible figures (27 figures in all). The choice figures are identified by letters and numbers (for example

a large red circle is R1) and the subject is required to print this identifying number on the answer sheet. The items are printed larger than those of most tests.

The MAC is essentially a power test with ample time limits which allow all but the very slow individual to finish all items. In order to finalize instructions to testees and check the time limits, a tryout was conducted by MacArthur with a small sample of Metis children at Slave Lake, Alberta.

The MAC appears to meet many of the criteria for tests designed to measure intellectual potential as outlined in the last chapter. The directions for administering the test call for an extensive teaching period and the level of language required for understanding of the directions is low. The extensive teaching period should help to familiarize the testee with specific instances as advocated by Campbell (1963) in the previously cited article. Gestures and demonstrations are used extensively in instructions during the teaching period and individual attention is given to all testees who are having difficulty. An effort is made to have every testee understand the nature of the problems to be solved and the method of finding and recording responses.

The MAC thus appears to meet some of the criteria for measures of intellectual potential but some of the criteria require statistical analysis. In this study, the criteria to be used will be numbers 1, 2, 3, 4, and 7 of the nine listed by MacArthur. Criteria numbers

5 and 9, dealing with stability and predictive validity are beyond the scope of this study. The MAC appears to meet MacArthur's criteria numbers 6 and 8 at least as well as do conventional group intelligence tests.

CHAPTER IV

DEFINITIONS, THEORY AND HYPOTHESES

I. Definitions

General Intellectual Ability - the ability to carry out tasks of a cognitive nature, this ability being the result of interaction between inherited ability and environmental factors. For the purposes of this study, general intellectual ability will be defined as the ability measured by tests having high loadings on the first factor extracted by factor analyzing a battery of tests of intellectual ability by the principal axes method (general intellectual ability factor).

Intellectual Potential - present potential for the development of intellectual ability. For the purposes of this study, intellectual potential will be defined as the ability measured by tests of intellectual ability, consisting of stimuli that are non-verbal in nature, and which have less bias against the members of cultural groups differing from the middle-class North American culture than do conventional group intelligence tests. Such tests should have high loadings on the general intellectual ability factor as defined above and low or negligible loadings on other, more specific factors.

Current School Achievement - due to differences in the test batteries administered to the two main samples in this study, two

different operational definitions of current school achievement will be used:

1. For the Eskimo sample, current school achievement will be defined as normalized standard scores derived from the Vernon Arithmetic, English and Vocabulary tests.

2. For the Indian-Metis sample, current school achievement will be defined as normalized standard scores derived from the Reading, Arithmetic, and Language grade-placement scores on the California Achievement Tests, Primary, Elementary and Junior High Levels, Form X.

Eskimos - subjects judged by their teachers to be of Eskimo parentage.

Indian-Metis - subjects who, in the opinion of their teachers, are regarded by the community as having Indian blood in their veins.

White - subjects judged by their teachers to be solely of white parentage.

II. Theory Underlying This Study

For the purposes of this study, intellectual ability will be considered to be hierarchical in nature with general intellectual ability at the top of the hierarchy. The abilities further down the hierarchy are of a more specific nature and are related to specific tasks. In using this model of intelligence the assumption is made that there is a general ability running through the more specific abilities or, in other words, these specific abilities are correlated.

The general ability at the top of the hierarchy may be thought of as being similar to Spearman's g but, as suggested by MacArthur (1964), not solely innate.

West points out that

> The statistical factor g or general ability is likely to enter heavily into a wide variety of future intellectual performances and may therefore be interpreted as a measure of present potential.
>
> (West, 1962, p. 28)

That is, the concept of intellectual potential is related to the concept of g postulated by Spearman. However, certain instruments that can be considered to be measuring g may be highly biased against cultural groups other than those for which they were primarily designed. The type of test most likely to fall into this category is the conventional group intelligence test which implicitly assumes that the testee has the ability to understand written language. This type of test, while it may be related to a person's "present potential . . . for the development of intelligent behavior" (West, 1962, p. 27) is likely to be culturally biased. The culture-reduced test must then measure g but be less dependent upon specific learnings, such as the learning of the English language, than are conventional intelligence tests.

General intellectual ability, or g, can best be measured through the use of a wide variety of cognitive tasks, including those depending on verbal stimuli. Such tests are closely bound to the culture in which they were constructed, not only because of the language used, but also because of the type of problem likely to be used.

For example, the type of problem used in a North American test of arithmetic reasoning may involve units of measurement, denominations of money, and problem situations which may not be familiar in a European, South African, or Eskimo setting.

Intellectual potential should therefore be measured through the use of stimuli likely to be as familiar as possible in all the cultural settings for which the instrument is designed. The stimuli should be non-verbal in nature; they might consist of symbols which have been found to be common to most cultures. The symbols should not force the subjects to make fine discriminations between small differences in cues presented in a two-dimensional form on paper.

Since general intellectual ability can best be measured by sampling a wide variety of cognitive tasks, an appropriate definition of this ability can be given in terms of factor analysis. General intellectual ability can be defined operationally as the common ability measured by tests having high loadings on a general factor extracted by factor analysis of a battery of tests sampling different types of cognitive tasks.

One appropriate method of factor analysis for defining the general intellectual ability factor is the method of principal axes. In the principal axes solution, the first factor is defined such that a maximum amount of the common variance of the battery of tests is accounted for. In other words, the first factor's contribution to the communalities of the tests in the battery will be the maximum

possible by any factor solution. A second factor, independent of the first, with a maximum contribution to the remaining common variance is found next. This process continues until the solution includes as many factors as can be given meaning.

The first principal axes factor extracted from a battery consisting of a number of measures of cognitive behavior could thus be thought of as a general intellectual ability factor.

Since intellectual potential is related to the concept of general intellectual ability, tests of intellectual potential must be sampling the same cognitive behavior as are sampled by tests of general intellectual ability. Therefore tests which are designed to measure intellectual potential should have high loadings on the general intellectual ability factor as discussed above. In other words, one of the important criteria for a test of intellectual potential is a high loading on the general intellectual ability factor.

Another important criterion for such a test is that cultural bias be minimized. There are several ways of determining the amount of cultural bias in a test, relative to that of other tests, but no way of determining an absolute amount of bias. In all likelihood a test having absolutely no bias would be impossible to construct.

One method of determining cultural bias would be to test samples of subjects from two different cultures with the tests to be compared, and then calculate the correlation coefficients between these tests and ethnic status (coded, for example 1 for the members

of one group and 2 for the members of the other). Then tests of the significance of the difference between correlation coefficients could be made in order to determine whether or not one test is more highly related to ethnic status than is the other test.

A second possible method of investigating the relative amount of cultural bias in tests is to test the significance of the difference between means of cultural groups on the different tests. In order to make such tests of significance, the scores of the different samples and different tests used must be on the same scale. One method of accomplishing such a common scale is to normalize the scores for a particular group and then use the raw scores and corresponding transformed scores as a table with which to convert the raw scores of the second cultural group, to an equivalent set of scores.

Since a person with a high level of intellectual potential should be capable of relatively more success in most classroom or employment situations than a person with a lower level of intellectual potential, tests designed to measure such potential should be correlated with school achievement or job success. However, such tests should not be too highly correlated with school success since success in a school situation is likely to be related to the philosophy and aims of education in the specific culture. In other words, intellectual potential should be related to a general kind of success in learning situations but not highly related to success in specific kinds of learning.

III. Hypotheses

With the above theoretical background the following hypotheses have been formulated in order to test the MAC against certain of MacArthur's criteria:

Hypothesis I - The MAC2 will have high positive loadings on the first factor extracted by the principal axes method of factor analysis for all Eskimo and Indian-Metis sub-samples.

Hypothesis II - The MAC2 will have low or negligible loadings on all factors except the first factor in a principal axes solution for all Eskimo and Indian-Metis sub-samples.

Hypothesis III - The scores used in testing hypothesis IIIa to IIIg will be standard scores based on normalized standard scores of the White sub-samples of the same age group.

IIIa: The mean MAC2 scores of the two Eskimo sub-samples will be significantly higher than the mean Otis scores of the same sub-samples.

IIIb: The mean MAC2 scores of the two Indian-Metis sub-samples will be significantly higher than the mean CTMM Language scores of the same sub-samples.

IIIc: The mean Otis scores of the two White sub-samples will be significantly higher than the mean scores of the Eskimo sub-samples of the same age groups.

IIId: The mean CTMM Language scores of the two White sub-samples will be significantly higher than the mean scores of the Indian-Metis sub-samples of the same age groups.

IIIe: There will be no significant differences between the mean MAC2 scores of the Eskimo sub-samples and the White sub-samples of the same age groups.

IIIf: There will be no significant differences between the mean MAC2 scores of the Indian-Metis sub-samples and the White sub-samples of the same age groups.

IIIg: There will be no significant differences between the mean MAC2 scores of the Eskimo and Indian-Metis sub-samples of the same age groups.

Hypothesis IV - Correlation coefficients calculated between the MAC2 and the measures of current school achievement will be moderate for all Eskimo and Indian-Metis sub-samples. A coefficient will be considered to be moderate if it is significantly greater than zero and lies between .300 and .600.

Validity, Reliability, and Item Analysis - In addition to the testing of the above hypotheses, certain estimates of the validity and reliability of the MAC2 will be calculated, and some item analysis procedures will be carried out as follows:

1. Split-half reliability: Pearson product-moment correlation coefficients between halves of the test, one-half consisting of the odd, and the other half of the even numbered items of the MAC2, will be calculated. The Spearman-Brown formula for correcting the coefficients for test length will be applied.

2. Kuder-Richardson formula 20 correlation coefficients will

be calculated for all sub-samples as further estimates of the reliability of the MAC2.

3. Point biserial correlation coefficients will be calculated between each item of the MAC2 and the total test scores.

4. The proportion of subjects passing each item will be calculated as an estimate of the difficulty level of the item.

5. Pearson product-moment correlation coefficients will be calculated between the MAC2 and other tests previously found to be culture-reduced. These will be used as estimates of concurrent validity of the MAC.

6. The results of the factor analyses and investigations of cultural bias will give an indication of construct validity for the MAC.

CHAPTER V

EXPERIMENTAL DESIGN

I. The Samples

During the spring of 1965, R. S. MacArthur of the University of Alberta conducted intellectual ability surveys among Indians and Metis at Faust, Alberta and Eskimos at Inuvik and Tuktoyaktuk, Northwest Territories. All children in grades three to eight attending selected schools in the survey areas were tested as well as some of the older children in grades one and two. Large batteries of tests including those listed below were administered to these two samples with the writer participating in the Faust test administration.

In addition to the two samples mentioned above, a sample of white pupils in grades three to eight attending a school in the city of Edmonton, was tested by the writer in the fall of 1965.

Each of the samples has been divided into two sub-samples on the basis of age, the younger group consisting of those subjects between the ages of 9 and 12 years and the older group of those between the ages of 12½ and 15½ years.

The Eskimo sample consisted of all the Eskimo children of both sexes between the ages of 9 and 15½ years enrolled in grades three to eight in the schools at Tuktoyaktuk and Inuvik (Protestant wing only). Also included were 21 grade one and two boys of age nine years eleven months or older. Missing data forced the dropping of 16 cases from

the sample. The outline of number of subjects by grade, sex and occupational level of parents for each sub-sample is given in Table I. Of the 167 Eskimos included in the sample, 77 were living in hostels while attending school and 88 were living at home. The other two were listed as living elsewhere. One hundred and fifty-five subjects reported that languages other than English were used in their homes, 89 reporting that the other language was used a small part of the time, 50 most of the time, and 16 all of the time. The other language was Eskimo for 153 of these subjects and Loucheux for the other two. The occupational levels of the parents of the Eskimo children involved of this study were rather low, with a mean Blishen Index score of 38.1 as compared to the Canadian average of about 50.

The Indian-Metis sample consisted of all the Indian and Metis children of both sexes attending school in grades three to eight in Faust at the time of the survey as well as eight children in the lower grades who were at least 9 years old. Faust is located about 200 miles north-west of Edmonton and has a population of about 1000. Many of the villagers are Indians or Metis of a very low socio-economic status. Indian and Metis families in the Faust area are very unstable because of the prevailing custom of common law marriages of short duration. The occupational levels of the parents of the Faust children involved in this study were about the same as those of the Eskimo sample, the mean Blishen Index score being 38.3. The outline

of number of subjects by grade, sex and occupational level of parents for each sub-sample can be seen in Table I. Of the 110 subjects in the Indian-Metis sample, 96 reported that Cree was spoken in the home, one reported French and one German. Fifty-three reported that the non-English language was spoken a small part of the time, 38 most of the time, and 7 all of the time. Ninety reported that they were living with at least one of their parents. Because of missing data 3 cases had to be dropped from the sample.

The White sample consisted of all of the white children of both sexes attending school in grades three to eight of a school in an older district of the city of Edmonton. Three of the children were dropped from the sample because they were recent immigrants and spoke very little English. The mean Blishen Index scores of the parent's occupations for the White sample was 50.7 which is very close to the Canadian average. The outline of number of subjects by grade, sex and occupational level of parents for each sub-sample can be seen in Table I. Of the 155 subjects in the White sample, 113 reported that only English was spoken in the home. Other languages reported were French, German, Dutch, Polish, Hungarian, Italian, Portuguese and Ukranian. Twenty-three reported that a non-English language was spoken in the home a small part of the time, 15 most of the time and 4 all of the time. Only two subjects were reported to be living with someone other than their natural parents and one of these two had been adopted as a baby.

TABLE I

NUMBER OF SUBJECTS BY GRADE SEX, AND OCCUPATIONAL LEVEL OF PARENTS FOR EACH SUB-SAMPLE

	Eskimo		Indian-Metis		White	
	9-12	12½-15½	9-12	12½-15½	9-12	12½-15½
Grade						
8		8		12		25
7		14		16	5	20
6	2	20	4	18	20	2
5	13	12	15	6	33	
4	19	15	18	0	24	
3	38	10	11	2	26	
2	10	1	8			
1	5					
Sex						
F	27	43	31	26	50	22
M	60	37	25	28	58	25
Oc. Lev.						
80-89					1	
70-79					1	
60-69		1			4	3
50-59	7	6	3	1	49	20
40-49	39	35	20	18	53	24
30-39	41	38	33	35		
Total N	87	80	56	54	108	47

Table II shows the number of subjects in each of the sub-samples as well as the mean age, in years, of each sub-sample.

TABLE II

SAMPLE SIZES AND MEAN AGES

Sample	9-12 years		12½-15½ years	
	N	Mean Age	N	Mean Age
Eskimo	87	10.8	80	13.7
Indian-Metis	56	10.6	54	13.8
White	108	10.0	47	13.2

II. The Variables

The variables used in this study are listed below along with the reasons that they were chosen and evidence of validity and reliability where the information was available. Unless otherwise indicated the variable was used with both the Eskimo and Indian-Metis samples. Only three variables were used with the White sample, the Otis Beta, the California Short Form Test of Mental Maturity and the MAC.

In selecting the variables to be used in the factor analyses, an attempt was made to include several kinds of tests of mental ability and also to include tests which had previously been found to be useful for cross-cultural comparisons of intellectual ability.

Some of the kinds of tests desired were: tests of intellectual ability that have been shown to be culturally-reduced, conventional verbal intelligence tests, and achievement tests yielding arithmetic and language scores. In order that a variety of cognitive tasks could be sampled, tests were selected that varied in the nature of the task presented to the testee. Thus, the first principal axes factor could be considered to be a general intellectual ability factor. Besides the tests, other variables included in the factor analyses were: age, grade, time in school, sex, parent's occupation level, and amount of non-English language used in the home. These variables were included in order to see if the MAC had group factor loadings in common with any of them.

The variables were:

1. Standard Progressive Matrices -- this test was chosen primarily because evidence suggests that it is one of the least bad culture-reduced tests of intellectual ability.

MacArthur (1962a, 1962b, 1964) reported first principal axes factor loadings ranging from .46 to .86 for samples of Indians and Metis from villages in Northern Alberta and the Northwest Territories. He also reported loadings of .60 and .42 for two samples of Rhodesian Africans. Elley (1961) reported a first principal axes loading of .71 for a sample of 271 grade 7 students in urban Alberta schools. These studies have also shown that the Progressive Matrices has few loadings on factors other than the general factor.

West (1962) in a study of bias in instruments measuring intellectual potential found that the Progressive Matrices was one of the best instruments for use with grades one through eight.

Elley reported that the Progressive Matrices has little correlation with ethnic status and Olson (1962) reported little correlation with language background.

Reports quoted in the test manual show re-test reliabilities varying with age from .83 to .93 and a correlation of .86 with the Terman-Binet. MacArthur (1964) reported correlations with the total score of the California Achievement Tests of from .26 to .55 for Metis in grades one to nine.

In administering the Progressive Matrices, the Coloured Progressive Matrices were used to teach the subjects how to solve the problems presented by the test items.

2. Safran Culturally Reduced Intelligence Test (SCRIT) -- evidence suggests that the SCRIT like the Progressive Matrices is one of the least bad culture-reduced measures of intellectual ability.

MacArthur reported first principal axes factor loadings ranging from .55 to .92 for the SCRIT. He also reported correlations with total California Achievement Test scores of from .35 to .49.

West found that the SCRIT met most of his criteria for culture-reduced instruments for measuring intellectual potential at

the levels for which it was thought to be appropriate, grades one to six.

In administering the SCRIT items 1 through 6 were used as teaching items so that with the four items given as practice items, a total of ten teaching items were used.

3. Lorge-Thorndike Intelligence Tests - Nonverbal, Level Three, Form A -- this test was chosen because it is made up of three sub-tests which use different kinds of items and these sub-tests have been shown to meet some of the criteria for culture-reduced instruments for measuring intellectual ability.

The Lorge-Thorndike total score had a very high g loading in Elley's factor analysis and MacArthur reported g loadings (first principal factor) of from .29 to .83 for the sub-tests, with most loadings falling between .60 and .80. He also reported g loadings ranging from .59 to .86 for the total test score.

West found that sub-test 2 met his criteria for three of the four grade levels, sub-test 3 for two levels and sub-test 1 for one grade level.

For the standardization sample, the split-half reliability was .93 and the parallel forms reliability .78. MacArthur found that correlations with the California Achievement Test total score ranged from .54 to .75 for grades one through nine.

4. Institute for Personality and Ability Testing Test of g, Scale two, Form A. (Cattell) -- this test was also selected because

evidence suggests that it is culture-reduced.

MacArthur reported g loadings for the Cattell ranging from .55 to .79. Elley found that the Cattell not only had the highest g loading in his battery but also had a very low correlation with socio-economic status. Olson reported a low correlation with language background.

Correlations with total scores on the California Achievement Tests reported by MacArthur were .41 and .38 for Metis children in grades 5-6 and 7-9 respectively.

5. California Short-Form Test of Mental Maturity, Elementary Level, Form S (CTMM) -- this test was included for two reasons. The Nonlanguage score yielded by this test has been shown to be less biased culturally than conventional group intelligence tests, and the language score can be considered to be a conventional, highly verbal measure of intelligence.

MacArthur found that the Nonlanguage scores had g loadings ranging from .32 to .88 and Elley reported a g loading of .62. MacArthur also found that the Nonlanguage scores correlated from .30 to .69 with the total score of the California Achievement Tests.

The publisher reported split-half reliability coefficients of .95, .91, and .95 respectively for the language, nonlanguage and total scores and correlations of .31 for the language scores with Otis Beta scores. Correlations of around .88 were reported between the CTMM total and the Stanford Binet.

The CTMM was used with the Indian-Metis and White samples only.

6. Otis Quick-Scoring Mental Ability Tests - Beta Test, Form EM -- this test was included as an example of the conventional group intelligence tests, likely to be culturally biased.

The publisher reported parallel forms reliability coefficients of between .69 and .98 for various samples and split-half reliability coefficients ranging from .89 to .94 for subjects in grades 4 to 8.

The Otis was used with the Eskimo and White samples only.

7. California Achievement Tests, Primary, Elementary and Junior High Levels, Form X -- this test was used in order to include certain measures of achievement in the battery. The scores used were the grade placements in Reading, Arithmetic and Language.

MacArthur reported g loadings ranging from .41 to .54 for these scores and Elley reported g loadings of from .36 to .44.

The California Achievement Tests were used with the Indian-Metis sample only. The primary level was used with grades 3 and 4, the elementary with 5 and 6 and the junior high with grades 7 and 8.

8. Vernon Arithmetic, English and Vocabulary Tests -- included in the battery as measures of achievement. The arithmetic test was Vernon's Graded Arithmetic-Mathematics test, the English an English attainment test constructed for use in the West Indies, and the Vocabulary the Mill-Hill Junior Form A.

Vernon (1965b) reported that these tests have high g loadings

(centroid) for samples of boys of the 11+ age in the West Indies and England. These tests also had loadings on an educational factor for the English sample and a verbal-educational factor for the West Indian sample.

These tests were used with the Eskimo sample only.

9. Other Tests from Vernon's Battery:

(a) Memory for Words -- this test was included as another example of a test likely to be culturally biased. The testee is given a list of twenty verbs which he studies for five minutes before being asked to write out as many of them as he can remember. Another two minutes study is then given and the subject is again asked to write out as many as he can remember. The test is scored one mark for each word regardless of spelling plus one-half a mark for each one in correct rank order. The total possible score is 60.

Vernon reported that this test had loadings on both general and educational factors for the English sample and on the general and verbal-cultural factors for the West Indian sample.

The test was used with the Eskimo sample only.

(b) Oral Information Memory Test -- this test was also considered to be a culturally biased instrument.

In the administration of the Oral Information Test, the tester slowly reads out a list of fifteen facts, emphasizing certain key words. Questions are subsequently asked on each fact. The questions are presented in a different order than are the facts. A

second period of reading the facts and asking questions follows the first and the scores of the two trials are added in order to form a composite score.

High loadings on the g factor are reported by Vernon for both the English and the West Indian samples as well as an educational loading for the former and verbal-educational loading for the latter.

(d) Gottschaldt Embedded Figures Test -- included as another example of a culture-reduced test of intellectual ability. The task given the testee is to find a figure embedded in a larger figure and trace it with a coloured pencil.

Vernon reported that the Embedded Figures had high g loadings for both the English and the West Indian samples and no educational or verbal-educational loadings.

10. MAC

11. Other Variables included in the Eskimo and Indian-Metis batteries were:

(a) age -- age in months

(b) grade -- grade in school

(c) total time in school -- number of years

(d) occupation of parent -- Blishen Index score for occupation (see Elley, 1961, Appendix B)

(e) sex -- 1 for females, 2 for males

(f) plans for age 20 -- first digit only of Blishen Index score for occupation chosen. This variable was used with the Eskimo battery only.

(g) Amount of language other than English spoken in the home (coded 1 for none, 2 for small part of the time, 3 for most of the time, and 4 for all of the time). This variable was used with the Indian-Metis sample only.

The order in which tests were administered was chosen such that between-sample differences in test-order effects were minimized. The tests were administered to the Indian-Metis sample in the same order as that used with the Eskimo sample. When different tests were used with the different samples, an attempt was made to administer similar tests at the same time in the testing schedule. The three tests used with the White sample were also administered in the same order as was used with the Eskimo and Indian-Metis samples.

III. Statistical Analysis

Hypotheses I and II -- The variables listed above were used in the test of these hypotheses. Separate analyses were made for each of the Eskimo and Indian-Metis sub-samples. The distributions of test scores were normalized and transformed so that the means were 50 and the standard deviations 10. The other variables were left in raw score form. Intercorrelations among scores were then calculated and the resulting correlation matrices were factor analyzed using the principal axes method as outlined by Harman (1960, pp. 154-191). In all cases, the number of factors extracted was determined by the number of eigenvalues found to be greater than one, as suggested by Kaiser (Harman, 1960, p. 363).

Hypothesis III -- The distribution of scores of the White sub-samples on the CTMM (language scores), Otis and MAC2 were normalized and transformed so that the means were 50 and the standard deviations 10. The resulting tables were then used to transform the raw scores of the Eskimo and Indian-Metis sub-samples on the same tests so that comparisons could be made between ethnic groups of the same age, and between different tests for the same sub-samples. Tests of the significance of the differences between the following paris of means were then made:

(a) MAC2 and Otis means of the two Eskimo sub-samples.

(b) MAC2 and CTMM Language means of the two Indian-Metis sub-samples.

(c) Otis means of White and Eskimo sub-samples of the same age groups.

(d) CTMM Language means of White and Eskimo sub-samples of the same age groups.

(e) MAC2 means of White and Eskimo sub-samples of the same age groups.

(f) MAC2 means of White and Indian-Metis sub-samples of the same age groups.

(g) MAC2 means of Eskimo and Indian-Metis sub-samples of the same age groups.

In each case, a test of the significance of the difference between variances was made in order to determine whether or not the assumption of homogeneity of variance was tenable. Where the assumption

was not tenable, Welch's approximation to Student's t distribution was used in the place of the usual formula for the calculation of t. The method used was that outlined by Winer (1962, pp. 37-39).

Hypothesis IV -- Pearson product-moment correlation coefficients were calculated between the MAC2 and the California Reading, Arithmetic, and Language tests for the two Indian-Metis sub-samples. Similarly, correlation coefficients were calculated between the MAC2 and the Vernon Arithmetic, English and Vocabulary tests for the two Eskimo sub-samples. Significance levels were found for all correlation coefficients by using a table of critical values of correlation coefficients (Ferguson, 1959, Table F, p. 315).

Validity, Reliability, and Item Analysis -- For each of the six sub-samples, the following estimates of the internal consistency of the MAC2 were calculated:

1. Pearson product-moment correlation coefficients between halves of the MAC2, made up by dividing the test into odd and even numbered items. The Spearman-Brown correction formula for estimating the reliability of the whole test from half a test was used to adjust the coefficients.

2. Kuder-Richardson Formula 20 reliability coefficients.

3. Point Biserial correlation coefficients between each item and the total test scores.

In addition to the above estimates of internal consistency, Pearson product-moment correlation coefficients with other tests that have been used to measure intellectual potential with minimum cultural

bias were calculated as estimates of concurrent validity for each of the Eskimo and Indian-Metis sub-sample. The other tests used were: Progressive Matrices, SCRIT, CTMM Nonlanguage (Indian-Metis only), Cattell and Lorge-Thorndike.

As a measure of the difficulty level of each item of the MAC2, the proportion of correct responses was calculated for each item for each sub-sample.

CHAPTER VI

RESULTS AND INTERPRETATIONS

Hypothesis I -- The MAC2 will have high positive loadings on the first factor extracted by the principal axes method of factor analysis for all Eskimo and Indian-Metis sub-samples.

The results of the factor analysis are presented in Tables III, IV, V and VI. The correlation matrices used in the factor analyses may be found in Appendix B. Hypothesis I appears to be supported for the data, the MAC2 loadings ranging from .489 to .619 on the general intellectual ability factor. In addition, the MAC2 loadings compare favorably with the loadings of other tests that have been previously identified as being less culturally biased than are conventional group intelligence tests. The range of general intellectual ability factor loadings of these other tests is from .478 to .874.

However, the first factor loadings for the age 12½ to 15½ Eskimos is lower than would be expected, and the loadings of other tests that have previously been identified as useful for measuring intellectual ability in a variety of cultures are generally higher than those of the MAC2 for all samples. It should be kept in mind when interpreting the results of the factor analyses that the communality of a test cannot exceed its reliability and that the reliability varies directly as the test length. If the MAC2 were increased in length its reliability should be increased and its communality may increase as a result. Thus the loadings on common factors such as the g factor would be higher.

Hypothesis II -- The MAC2 will have negligible or low loadings on all factors except the first factor in a principal axes solution for all Eskimo and Indian-Metis sub-samples.

As can be seen by Tables III, IV, V and VI, this hypothesis is supported for three of the four sub-samples, the MAC2 having only one small loading on factors other than the general intellectual ability factor for two sub-samples and no loadings for a third. However, the data of Table IV (Eskimos 12½ - 15½) do not support the hypothesis, the MAC2 having loadings on three factors besides that on the general intellectual ability factor.

The MAC2 seems to compare favorably with other culture reduced tests of intellectual ability such as the Progressive Matrices, Cattell, and SCRIT, with respect to the number of loadings on factors other than the general factor.

Hypothesis IIIa -- The mean MAC2 scores of the two Eskimo sub-samples will be significantly higher than the mean Otis scores of the same sub-samples.

This hypothesis was supported for the age 12½ to 15½ sub-sample but not for the age 9 to 12 sub-sample. For the older group the difference was significant at the .001 level. These results suggest that the MAC is less culturally biased than is the Otis for Eskimos of junior high school age but not for younger Eskimos.

Hypothesis IIIb -- The mean MAC2 scores of the two Indian-Metis sub-samples will be significantly higher than the mean CTMM Language scores of the same sub-samples.

This hypothesis was supported for the age 12½ to 15½ sub-sample

TABLE III

ESKIMOS 9-12 YEARS UNROTATED PRINCIPAL

AXES FACTOR LOADINGS

Variable	Factor Loadings					Commu-nality
	I	II	III	IV	V	
1. Grade	.888	.302				.909
2. Age	.469		-.549		-.463	.769
3. Time-school	.635					.567
4. Occ.-parent		.390		.584	.520	.839
5. Plans-20	.313			.664		.659
6. Sex		-.484	-.661		.337	.814
7. P.M.	.679	-.589				.809
8. SCRIT	.571	-.355	.306			.553
9. MAC2	.609				.337	.639
10. L-Th 1	.673					.530
11. L-Th 2	.838					.775
12. L-Th 3	.751	-.329				.704
13. Cattell	.734	-.430				.731
14. Otis	.743					.666
15. V. Arith.	.909					.893
16. Mem. Wds.	.578	.363			-.317	.640
17. V. Eng.	.866					.823
18. V. Vocab.	.798	.318				.748
19. Let. & No.	.807					.701
20. Oral Info.	.790					.702
21. Emb. Fig.	.684	-.368				.630
Percentage of Total Variance	46.8	9.9	5.6	5.1	4.6	71.9

Note: Loadings between -.300 and .300 ommitted.

TABLE IV

ESKIMOS 12½-15½ YEARS UNROTATED PRINCIPAL AXES FACTOR LOADINGS

Variable	Factor Loadings					Communality
	I	II	III	IV	V	
1. Grade	.873	.375				.927
2. Age	.366			.587		.667
3. Time-school	.707	.474				.770
4. Occ.-parent			.457		-.729	.780
5. Plans-20	.371		-.457		-.483	.618
6. Sex		-.329	-.504	.641		.789
7. P. M.	.798					.721
8. SCRIT	.520	-.487	.349			.709
9. MAC2	.489	-.313	.453		.334	.654
10. L-Th 1	.606			.363		.560
11. L-Th 2	.826					.734
12. L-Th 3	.680	-.417				.652
13. Cattell	.478	-.643				.668
14. Otis	.855					.808
15. V. Arith.	.894					.835
16. Mem. Wds.	.414				-.331	.352
17. V. Eng.	.871	.300				.871
18. V. Vocab.	.729	.373				.714
19. Let. & No.	.843					.735
20. Oral Info.	.715					.623
21. Emb. Fig.	.755					.654
Percentage of Total Variance	43.9	9.9	6.3	5.5	5.2	70.7

Note: Loadings between -.300 and .300 ommitted.

TABLE V

INDIAN-METIS 9-12 YEARS UNROTATED PRINCIPAL AXES FACTOR LOADINGS

Variable	Factor Loadings					Communality
	I	II	III	IV	V	
1. Grade	.848	-.449				.927
2. Sex				-.692	-.567	.896
3. Age	.511	.601	.362			.777
4. Time-school	.719	-.591				.916
5. Occ.-parent			-.787			.767
6. Lang-Home	-.402		.580	.382		.722
7. P. M.	.667	.310	.441			.792
8. SCRIT	.637			.370		.669
9. MAC2	.619					.457
10. L-Th 1	.515	.512		-.411		.474
11. L-Th 2	.721	.375				.745
12. L-Th 3	.612	.485			.410	.787
13. Cattell	.618		.307			.579
14. Cal. Rdg.	.915					.868
15. Cal. Arith.	.937					.913
16. Cal. Lang.	.862					.813
17. Let. & No.	.805	.320				.816
18. Oral Info.	.889					.869
19. Emb. Fig.	.725					.607
20. CTMM Lang.	.812					.672
21. CTMM Non-L.	.698				-.511	.762
Percentage of Total Variance	48.2	9.5	7.9	5.8	5.2	76.7

Note: Loadings between -.300 and .300 omitted.

TABLE VI

INDIAN-METIS 12½-15½ YEARS UNROTATED PRINCIPAL AXES FACTOR LOADINGS

Variables	Factor Loadings					Communality
	I	II	III	IV	V	
1. Grade	.635	.597				.825
2. Sex				.869		.836
3. Age		.749	-.302			.747
4. Time-school		.762				.667
5. Occ.-parent					.872	.804
6. Lang-home	-.469		.335			.426
7. P. M.	.874					.769
8. SCRIT	.833					.806
9. MAC2	.580		-.306			.556
10. L-Th 1	.618			.329		.539
11. L-Th 2	.766					.662
12. L-Th 3	.794					.762
13. Cattell	.730		-.399			.782
14. Cal. Rdg.	.814		.366			.853
15. Cal. Arith.	.769					.756
16. Cal. Lang.	.780		.408			.847
17. Let & No.	.874					.824
18. Oral Info.	.440		.624			.675
19. Emb. Fig.	.816		.361			.624
20. CTMM Lang.	.816		.361			.807
21. CTMM Non-L.	.592			.333		.495
Percentage of Total Variance	42.9	8.5	7.8	6.8	5.7	71.7

Note: Loadings between -.300 and .300 omitted.

but not for the age 9 to 12 sub-sample. For the older group the difference was significant at the .001 level but for the younger group there was a significant difference at the .05 level favoring the CTMM. As with the Eskimo sample, the results suggest that the MAC is culture-reduced for junior high school aged children but not for the younger school children.

The data showing the results of the tests of hypotheses IIIa and IIIb are presented in Table VII.

Hypothesis IIIc -- The mean Otis scores of the two White sub-samples will be significantly higher than the mean scores of the Eskimo sub-samples of the same age groups.

The data support this hypothesis, the differences being significant at the .005 level for both sub-samples tested against the White sub-samples. Thus, support is given the thesis that the Otis Beta is culturally biased.

Hypothesis IIId -- The mean CTMM Language scores of the two White sub-samples will be significantly higher than the mean scores of the Indian-Metis sub-samples of the same age group.

The data support this hypothesis, the differences being significant at the .005 level for both sub-samples tested against the White sub-samples. Thus support is given the thesis that the CTMM Language scores are culturally biased.

Hypothesis IIIe -- There will be no significant differences between the mean MAC2 scores of the Eskimo sub-samples and the White sub-samples of the same age groups.

The data support this hypothesis for the 12½ to 15½ year old sub-sample, there being no significant difference at the .05 level. For the 9 to 12 year old sub-sample, however, the data do not support the hypothesis, the difference being significant at the .005 level. The data thus support the findings under hypothesis IIIa.

Hypothesis IIIf -- There will be no significant differences between the mean MAC2 scores of the Indian-Metis sub-samples and the White sub-samples of the same age groups.

The data do not support this hypothesis. For the 9 to 12 year old sub-sample the difference is significant at the .005 level and for the 12½ to 15½ year old sub-sample the difference is significant at the .01 level.

Hypothesis IIIg -- There will be no significant difference between the mean MAC2 scores of the Eskimo and Indian-Metis sub-samples of the same age groups.

For the age 9 to 12 sub-samples the difference was significant at the .05 level, with the Eskimo mean being larger. However, for the 12½ to 15½ year old sub-samples the difference was not significant at the .05 level. There were no significant differences between the variances of the Eskimo and Indian-Metis scores on the MAC2 for either age group.

In general, the various parts of hypothesis III were supported by the data for the age 12½ to 15½ Eskimos and Indian-Metis but not for the younger sub-samples. Differences between variances were

found to be significant in all cases on the conventional tests while there were no significant differences between variances on the MAC2. The data for the tests of hypotheses IIIa to IIIg are presented in Tables VII, VIII and IX.

TABLE VII

DIFFERENCES BETWEEN MAC2 AND CONVENTIONAL TESTS

Sub-Sample	MAC2 Mean	Conventional Test and Mean		t	Significance Level
Esk 9-12	44.52	Otis	44.51	0.01	ns
Esk 12½-15½	47.59	Otis	31.79	13.84	.001
I-M 9-12	40.52	CTMM-L	44.06	2.65	.05
I-M 12½-15½	45.52	CTMM-L	34.37	8.50	.001

As an additional indication of relative cultural bias, correlation coefficients between the tests involved and ethnic status were calculated. In the calculation of these correlations ethnic status was coded 1 for Whites, 2 for Indian-Metis and 3 for Eskimos. Therefore a negative correlation coefficient indicates that the test favors the white sample. All coefficients are significantly different from zero except those for the 12½ to 15½ year old groups on the MAC2. The correlation coefficients are presented in

TABLE VIII

DIFFERENCES BETWEEN ETHNIC GROUP MEANS

Sub-Sample	Test	Eskimo Mean	White Mean	t	Sig. Level
9-12	Otis[a]	44.5	50.0	4.44	.005
9-12	MAC2	44.5	50.3	3.71	.005
12½-15½	Otis[a]	31.8	50.0	11.27	.005
12½-15½	MAC2	47.6	50.4	1.40	ns

Sub-Sample	Test	Ind-Met. Mean	White Mean	t	Sig. Level
9-12	CTMM-L[a]	44.1	49.8	4.13	.005
9-12	MAC2	40.5	50.3	5.94	.005
12½-15½	CTMM-L[a]	34.4	50.2	9.22	.005
12½-15½	MAC2	45.5	50.4	2.49	.01

Sub-Sample	Test	Eskimo Mean	Ind-Met. Mean	t	Sig. Level
9-12	MAC2	44.5	40.5	2.16	.05
12½-15½	MAC2	47.6	45.5	1.23	ns

[a]Welch's approximation to t used because variances differed significantly from each other.

TABLE IX

DIFFERENCES BETWEEN ETHNIC GROUP VARIANCES

Sub-Sample	Test	Eskimo Var.	White Var.	F	Sig. Level
9-12	Otis	54.2	98.0	1.81	.05
9-12	MAC2	137.4	95.5	1.08	ns
12½-15½	Otis	46.4	96.0	2.07	.05
12½-15½	MAC2	123.7	97.2	1.27	ns

Sub-Sample	Test	Ind-Met. Var.	White Var.	F	Sig. Level
9-12	CTMM-L	56.3	98.0	1.74	.05
9-12	MAC2	102.6	95.5	1.08	ns
12½-15½	CTMM-L	46.7	98.0	2.10	.05
12½-15½	MAC2	90.1	97.2	1.08	ns

Sub-Sample	Test	Eskimo Var.	Ind-Met. Var.	F	Sig. Level
9-12	MAC2	137.4	102.6	1.34	ns
12½-15½	MAC2	123.7	90.1	1.37	ns

Table X along with the results of tests of the significance of differences between MAC2 and conventional test correlations with ethnic status.

TABLE X

CORRELATION COEFFICIENTS WITH ETHNIC STATUS

Sub-Samples	Test	r	t	p Level
Eskimos and Whites 9-12	Otis MAC2	-.322 -.286	.501	ns
Eskimos and White 12½-15½	Otis MAC2	-.717 -.082	8.490	.001
Ind-Met. and Whites 9-12	CTMM-L MAC2	-.310 -.426	1.473	ns
Ind-Met. and Whites 12½-15½	CTMM-L MAC2	-.528 -.149	3.464	.01

The correlational data support the previous findings that the MAC2 is less biased than are the conventional tests for the age 12½ to 15½ groups but not for other groups.

Hypothesis IV -- Correlation coefficients calculated between the MAC2 and the measures of current school achievement will be moderate for all Eskimo and Indian-Metis sub-samples. A coefficient will be considered to be moderate if it is significantly greater than zero and lies between .300 and .600.

The data support this hypothesis for nine of the twelve correlation coefficients calculated. Table XI presents the data.

TABLE XI

CORRELATION COEFFICIENTS - MAC WITH ACHIEVEMENT

Sub-Samples	Test	r	Sig. Level
Eskimos 9-12	Arith.	.504	.01
	Eng.	.441	.01
	Vocab.	.379	.01
Eskimos 12½-15½	Arith.	.364	.01
	Eng.	.257	.02
	Vocab.	.142	ns
Indian-Metis 9-12	Arith.	.575	.01
	Lang.	.519	.01
	Rdg.	.532	.01
Indian-Metis 12½-15½	Arith.	.252	ns
	Lang.	.319	.02
	Rdg.	.369	.01

Validity, Reliability and Item Analysis -- The MAC appears to have construct validity for the age 12½-15½ sub-samples of Eskimos and Indian-Metis since it meets the criteria of high g loadings on the general intellectual ability factor and is less biased culturally than are conventional group tests of intelligence.

Concurrent validity of the MAC is shown by the correlation coefficients with other culture-reduced tests. Of the eighteen coefficients calculated, 16 are significantly different than zero

at the .01 level, one at the .05 level and only one is not significantly different than zero. The coefficients are shown in Table XII.

TABLE XII

MAC2 CORRELATIONS WITH OTHER

CULTURE-REDUCED TESTS

Sub-Sample	Prog. Mat.	SCRIT	Cattell	Lorge-Thorn.	CTMM Non-Lang.
Eskimos 9-12	.348**	.365**	.407**	.500**	
Eskimos 12½-15½	.404**	.284**	.253*	.346**	
Ind-Met. 9-12	.417**	.454**	.355**	.550**	.369**
Ind-Met. 12½-15½	.434**	.465**	.472**	.453**	.183

*Significant at .05 level.

**Significant at .01 level.

Evidence of internal consistency within the MAC2 is shown in Tables XIII and XIV, Table XIII showing high split-half correlation coefficients and Kuder-Richardson formula 20 coefficients. Table XIV shows high median point biserial correlation coefficients between most of the MAC2 items and the total test scores (median of 6 sub-samples).[1]

[1]Data for each sub-sample is presented in Appendix D.

The data of Table XIV also shows the median difficulty levels of each of the eighteen items of the MAC2 over the six sub-samples.

TABLE XIII

INTERNAL CONSISTENCY OF MAC2

Sub-Sample	Split-half r	Corrected Split-half	KR 20 r
Eskimos 9-12	.792	.884	.819
Eskimos 12½-15½	.775	.873	.793
Indian-Metis 9-12	.679	.809	.837
Indian-Metis 12½-15½	.437	.609	.631
Whites 9-12	.759	.863	.791
Whites 12½-15½	.735	.847	.813

TABLE XIV

CORRELATIONS WITH TOTAL SCORE AND ITEM

DIFFICULTY LEVELS

Item No.	Point Biserial r	Difficulty Level
1	.26	.96
2	.51	.89
3	.51	.92
4	.65	.94
5	.63	.82
6	.59	.89
7	.55	.88
8	.68	.79
9	.61	.93
10	.61	.70
11	.24	.39
12	.66	.72
13	.55	.57
14	.28	.10
15	.31	.08
16	.41	.12
17	.23	.04
18	.08	.01

Note: Figures are medians over 6 sub-samples.

CHAPTER VII

DISCUSSION AND IMPLICATIONS

I. Evaluation of the MAC - MacArthur's Criteria

The following section is devoted to an evaluation of the MAC as a culture-reduced instrument for the assessment of intellectual potential. The MAC will be evaluated against those of the previously mentioned nine criteria for such a test which were studied in this investigation.

1. It should largely sample the broad factor of general intellectual ability running through a wide variety of European-American kinds of intellectual tasks, including arithmetic, linguistic, and reading achievement.

The results of the factor analyses show that the MAC2 in its preliminary form meets this criterion, although the Progressive Matrices appears to have met the criterion better than did the MAC2. The SCRIT met this criterion about as well as did the MAC2. The median first principal factor loadings were .595, .739 and .604 respectively for the MAC2, Progressive Matrices and SCRIT, and the median communalities were .605, .781 and .689.

2. It should have negligible loadings on verbal, numerical, and other group factors. That is, it should sample general intellectual ability through simple basic symbols not very dependent upon particular previous learnings.

The MAC appears to have met this criterion as well as did the SCRIT but not as well as did the progressive Matrices. The MAC and the SCRIT each had a total of five loadings on group factors, while the Progressive Matrices had only three. As can be seen by Tables III, IV, V and VI, the sub-sample for which the MAC met this criterion least well was the Eskimo 12½ to 15½ year old group. Three of the five loadings that the MAC had on group factors were with this sub-sample.

3. It should show less difference between cultures than do alternate measures of intelligence.

The MAC met this criterion more adequately for the older sub-samples of Eskimos and Indian-Metis than it did for the younger sub-samples. In testing differences between White and Eskimo or Indian-Metis variances, it was found that, at the .05 level, there were significant differences between ethnic groups on the Otis and CTMM language but not on the MAC2. In testing differences between White and Eskimo or Indian-Metis means, it was found that there were no significant differences between Eskimos and Whites in the 12½ to 15½ year old range on the MAC2 but differences at the 9 to 12 year old range were significant. In addition, differences between Whites on the one hand, and Eskimos or Indians and Metis on the other, were significant at the .01 level on the conventional tests at both age levels. Thus, although the Otis and CTMM Language scores had very high loadings on the general intellectual ability factors, as can be seen by Tables III, IV, V and VI, they are also culturally biased.

It would be expected that this bias would result in unreliable scores for those Eskimo and Indian or Metis children who had most recently started school, and many of the children in the areas from which the samples of this study were drawn start school at older ages than do white children in an urban center.

The data of this study, therefore, support the hypothesis that the MAC in its present form, has less cultural bias than do the Otis Beta or the CTMM Language for Eskimo and Indian or Metis children respectively in the 12½ to 15½ year old range. However, for the younger Eskimos and Indians or Metis, the MAC could not be called culture-reduced, or less biased culturally than the ordinary verbal group test of intelligence.

4. It should show moderate relationship with current school achievement.

The data show that the MAC meets this criterion adequately for the four Eskimo and Indian-Metis sub-samples. Of the twelve correlation coefficients calculated, only two failed to reach significance and two more fell outside the range of from .300 to .600 defined as the range of moderate coefficients.

5. It should show evidence of stability under changed environmental conditions relative to that of alternative measures of intelligence.

This criterion was not investigated in this study.

6. It should minimize effects of test sophistication, providing

plenty of appropriate practice experience, having directions depending little if at all on language, and being unspeeded.

No empirical investigations were made with respect to this criterion but the MAC appears to meet it adequately. Included in the directions is an extensive teaching period and the directions rely heavily upon gestures and demonstrations. The time limits are set so that all but the very slow individuals are able to finish.

7. It should be reliable.

The median reliability coefficients across the six sub-samples, found for the MAC2 by the split-half method and by the use of the Kuder-Richardson formula 20, were .855 (corrected for test length) and .803 respectively.

8. It should be practical and usable from an administrative standpoint.

The MAC proved to be an easy test to administer and the extensive teaching period prior to the actual testing appeared to be adequate for acquainting the subjects with the nature of the task and method of recording responses. However, the separate choice sheet made the handling of the test material difficult for the younger children. The test booklets are rather large and leave little room for the choice sheet on a small desk.

The subjects appeared to find the test interesting and many of them said they enjoyed doing it.

9. When and if possible, it should show long-term validity as a predictor of success in intellectual tasks when appropriate

adaptive intervening treatments have been employed.

This criterion was not studied in this investigation.

II. General Discussion

The results of this study suggest that the MAC, in its present form, shows some promise as an instrument for the measurement of intellectual potential with a minimum of cultural bias, especially for children of junior high school age. Studies with ethnic groups other than Eskimos and Indians and Metis are necessary before any generalized use will be possible. Certain changes may prove fruitful if the test is to be used to any extent and it may be possible to use a revised form of the MAC with children of elementary school age.

The fact that it was the younger sub-samples of Eskimos and Indian-Metis who scored significantly lower than Whites of the same age group suggests that the MAC may be too difficult for them. This may be at least partly due to the format of the test which, as pointed out above may make it difficult for the younger children to manipulate. The White children of the younger ages have likely had more practice with manipulating papers and this may partially account for their being more successful on the MAC than were the Eskimos and Indian-Metis. Perhaps some of the difficulties of the test could be eliminated by changing the format. A format similar to that of the Progressive Matrices with one item per page and a multiple choice type of response, with the possible responses printed on the same page may be better.

Another change that may improve the MAC is the increasing of the number of items included. The MAC2 consisted of only 18 items and the reliability could likely be improved by increasing the length of the test. Some ready sources of additional items are the MAC1 and MAC3, but it may also be necessary to construct new items.

In discussing the difficulty level of test items, Remmers, Gage & Rummel (1960) state that in order to maximize the efficiency of measuring individual differences there should be a wide range in difficulty level, with a few very difficult items, a few very easy items and the remaining items should be evenly distributed along the range of difficulties.

As can be seen by Table XIV the median difficulty levels of the MAC2 items for the sub-samples of this study show that there are more items that might be classified as very easy or very difficult then there are items falling between these two extremes in difficulty. Of the eighteen items there were five that were correctly solved by fewer than 20 per cent of the subjects and eight items that were correctly solved by more than 80 per cent of the subjects. If the suggestions of Remmers et. al. are to be followed, the MAC2 should include more items in the range of difficulty level between .20 and .80. This should be kept in mind when new items are added to the MAC2. The items of the MAC1 and MAC3 should be investigated and those items having difficulty levels falling between .20 and .80 should be considered for inclusion in the MAC2.

A third possible way of improving the MAC2 would be to add items of a different nature than those presently included. One type of item that might be included is the figure analogy item similar to those of sub-test 3 of the Lorge-Thorndike. West (1962) found that this sub-test met his criteria for culture-reduced tests at two of the four grade levels studied and the factor analyses of the present study show high g loadings and very low loadings on group factors. This type of item could be constructed with the same symbols as are presently used in the MAC. Another possible type of item is the matrix item similar to that of sub-test 3 of the Cattell. The same simple symbols could also be used in the matrices.

In view of the extreme difficulty levels and the small spread in raw scores on the MAC2 (see Appendix A for raw score means and standard deviations), the results of this study suggest that the MAC shows promise as a culture-reduced instrument for measuring intellectual potential. Some modifications such as those suggested above may greatly improve the test, especially if a better distribution of difficulty levels and wider spread in scores results from these modifications.

III. Implications

The fact that it may be possible to measure intellectual potential with relatively little bias against persons from non-white-middle-class-European-American backgrounds has implications for the diagnosis and planning of adaptive treatments for immigrants, as well as for culturally-deprived native peoples. Toward the end of better educational

diagnosis, research should be carried out with the MAC with groups other than Eskimos and Indians. At the same time it must be recognized that individual prediction is difficult and that diagnosis with only the results of one test is not very reliable or valid.

The results of this study are encouraging in that they tend to support the hypothesis that tests of intellecutal ability can be developed that may be used in a variety of cultures. They also appear to support the hypothesis that the use of simple symbols that are likely to be familiar in a variety of cultures is a necessity in cross-cultural testing. The author feels that the MAC has shown promise as a test of intellectual potential with a minimum of cultural bias, at least for children of junior high school age. This promise may be fulfilled, and the test may also prove useful for younger children. Further research on the use of the MAC is necessary in order to evaluate the test more completely.

Some of the directions that may be taken in further research are those suggested above: use of the MAC with other cultural groups such as immigrants, and use of a modified form of the MAC with a different format so that a separate choice sheet is unnecessary and the number of possible responses is decreased. It would also be desirable to add more items to the MAC in order to increase the reliability and possible range of marks. The suggested modifications in the format may decrease the difficulty level of some of the harder items. Perhaps a necessary prerequisite to some of these suggested investigations

is a more extensive item analysis of the MAC.

The postulate that there is a general intellectual ability running through the more specific abilities measured by intelligence tests is supported by this study. Also supported is the thesis that the conventional, highly verbal intelligence test is more culturally biased than is a test such as the MAC which seems to be measuring intellectual ability through the use of nonverbal stimuli.

REFERENCES

REFERENCES

Anderson, W. F. Relation of Lorge-Thorndike Intelligence Test Scores of Public School Pupils to the Socio-economic Status of Their Parents. J. Exp. Educ., 1962, 31, 73-76.

Campbell, A. C. Intelligence Test Items and Serendipity. Australian J. Psychol., 1963, 15, 201-205.

Cattell, R. B., & Cattell, A. K. S. Handbook for the Individual or Group Culture Free Intelligence Test, Scale 2. Champagne, Illinois: Institute for Personality and Ability Testing, 1958.

Eells, K. W., Davis, A., Havighurst, R. J., Herrick, E., & Tyler, R. Intelligence and Cultural Differences. Chicago: Univer. of Chicago Press, 1951.

Elley, W. B. A Comparative Analysis of the Socio-economic Bias in Selected Intelligence Tests. Unpublished doctoral dissertation, Univer. of Alta., 1961.

Ferguson, G. A. Statistical Analysis in Psychology and Education. Toronto: McGraw-Hill, 1959.

Harman, H. H. Modern Factor Analysis. Chicago: Univer. of Chicago Press, 1960.

Hess, R. D. Controlling Culture Influence in Mental Testing - An Experimental Test. J. of Educ. Res., 1955, 49, 51-58.

Kidd, Aline H., & Rivoire, Jeanne L. The Culture-fair Aspects of the Development of Spatial Perception. J. Genet. Psychol., 1965, 106, 101-111.

Knapp, R. R. The Effects of Time Limits on the Intelligence Test Performance of Mexican and American Subjects. J. Educ. Psychol., 1960, 51, 14-20.

Knief, Lotus M., & Stroud, J. B. Intercorrelations Among Various Intelligence, Achievement, and Social Class Scores. J. Educ. Psychol., 1959, 50, 117-120.

Lorge, I., & Thorndike, R. L. Examiner's Manual, The Lorge-Thorndike Intelligence Tests. Boston: Houghton Mifflin, 1957.

MacArthur, R. S. Summary of a Tentative Rationale for Assessing Intellectual Potential with Minimum Cultural Bias. Paper read at University of London Institute of Education Senior Oversea Seminar, London, December, 1962. (a)

MacArthur, R. S. The Construct Validity of a Number of Measures of Intellectual Potential for an Alberta Metis Sample. Paper read at Canadian Psychol. Ass., Hamilton, 1962. (b)

MacArthur, R. S. Minimizing Cultural Bias in Assessing Intellectual Potential - Arctic and Rhodesian Samples. Paper read at Canadian Psychol. Ass., Halifax, June, 1964.

Marquart, Dorothy I., & Bailey, Lois L. An Evaluation of the Culture Free Test of Intelligence. J. Genet. Psychol., 1955, 86, 353-358.

Olson, D. R. The Influence of Foreign Language Background on Performance on Selected Intelligence Tests. Unpublished master's thesis, Univer. of Alta., 1962.

Otis, A. S. Otis Quick-Scoring Mental Ability Tests. Manual of Directions for Beta Test. Yonkers-on-Hudson: World Book, 1954.

Raven, J. C. Progressive Matrices. London: Lewis, 1947.

Raven, J. C. Standard Progressive Matrices. London: Lewis, 1956.

Remmers, H. H., Gage, N. L., & Rummel, J. F. A Practical Introduction to Measurement and Evaluation. New York: Harper and Row, 1960.

Schwarz, F. A. Adapting Tests to the Cultural Setting. Educ. Psychol. Measmt., 1963, 23, 673-686.

Sullivan, E. T., Clark, W. W., & Tiegs, E. W. Manual, California Short-Form Test of Mental Maturity (Elementary). Los Angeles: California Test Bureau, 1957.

Tiegs, E. W., & Clark, W. California Achievement Tests (Primary Level). Los Angeles: California Test Bureau, 1957.

Tiegs, E. W., & Clark, W. California Achievement Tests (Elementary Level). Los Angeles: California Test Bureau, 1957.

Tiegs, E. W., & Clark, W. California Achievement Tests (Junior High Level). Los Angeles: California Test Bureau, 1957.

Tyler, F. T. Comments on the Correlational Analysis Reported in Intelligence and Cultural Differences. J. Educ. Psychol., 1953, 44, 288-295.

Vernon, P. E. The Structure of Human Abilities. London: Methuen, 1956.

Vernon, P. E. Environmental Handicaps and Intellectual Development. Brit. J. Educ. Psychol., 1965, 35, 9-20. (a)

Vernon, P. E. Environmental Influences and Mental Development. Paper read at Univer. of Alta., Edmonton, March, 1965. (b)

Wert, F. E., Neidt, C. O., & Ahmann, J. S. Statistical Methods in Educational and Psychological Research. New York: Appleton Century Crofts, 1954.

West, L. W. Assessing Intellectual Ability with a Minimum of Cultural Bias for Two Samples of Metis and Indian Children. Unpublished master's thesis, Univer. of Alta., 1962.

West L. W. & MacArthur, R. S. An Evaluation of Selected Intelligence Tests for Two Samples of Metis and Indian Children. Alta. J. Educ. Res., 1964, 10, 17-27.

Yates, A. J. A Further Study of Progressive Matrices. Brit. J. Educ. Psychol., 1963, 33, 307-311.

APPENDICES

APPENDIX A

MAC - RAW SCORE MEANS AND

STANDARD DEVIATIONS

Sample	Mean	S.D.
Eskimos 9-12	8.83	3.58
Eskimos 12½-15½	11.55	3.02
Indian-Metis 9-12	7.68	3.66
Indian-Metis 12½-15½	11.31	2.14
Whites 9-12	10.69	2.64
Whites 12½-15½	12.04	2.67

APPENDIX B

ABBREVIATIONS USED IN

THE FOLLOWING TABLES

Gr. - grade in school
Age - age in months to nearest month
TIS - total time in school
Occ - occupation of parent, Blishen Index score
Pl. - plans for age 20, first digit only of Blishen Index score of occupation indicated
Sex - 1 for female, 2 for male
PM - Progressive Matrices normalized scores
SCR - Saffran Culturally Reduced Intelligence Test normalized scores
MAC - MAC normalized scores
LT1 - sub-test 1 of Lorge-Thorndike Intelligence Test - Nonverbal, normalized scores
LT2 - sub-test 2 of Lorge-Thorndike Intelligence Test - Nonverbal, normalized scores.
LT3 - sub-test 3 of Lorge-Thorndike Intelligence Test - Nonverbal, normalized scores
Cat - Institute for Personality and Ability Testing, Test of g (Cattell), normalized scores
Oti - Otis Beta normalized scores
Ari - Vernon Arithmetic Test normalized scores (Eskimos)
- California Achievement Tests, arithmetic grade-placement, normalized scores (Indian-Metis)
Eng - Vernon English Test normalized scores (Eskimos)
Voc - Vernon Vocabulary Test normalized scores (Eskimos)
L&N - Letters and Numbers Test normalized scores
OI - Oral Information Memory Test normalized scores
EF - Gottschaldt Embedded Figures Test normalized scores
Amt - amount of Language other than English used in the home (Indian-Metis)
Rdg - California Achievement Tests, Reading grade-placements, normalized scores (Indian-Metis)
Lan - California Achievement Tests, Language grade-placements, normalized scores (Indian-Metis)
CTL - California Short Form Test of Mental Maturity normalized language scores (Indian-Metis)
CTN - California Short Form Test of Mental Maturity normalized non-language scores (Indian-Metis)

INTERCORRELATIONS - FACTOR ANALYSIS VARIABLES - ESKIMOS 9-12 *

	Gr.	Age	TIS	Occ	Pl.	Sex	PM	SCR	MAC	LT1	LT2	LT3	Cat	Oti	Ari	MFW	Eng	Voc	L&N	OI	EF
Gr.	1.0																				
Age	444	1.0																			
TIS	667	365	1.0																		
Occ	274	014	156	1.0																	
Pl.	286	084	251	022	1.0																
Sex	-114	165	-036	-072	-015	1.0															
PM	434	398	264	-067	063	181	1.0														
SCR	386	236	237	017	106	-045	582	1.0													
MAC	482	060	441	180	240	-033	381	367	1.0												
LT1	503	301	348	180	195	-019	480	424	282	1.0											
LT2	697	272	499	012	274	-020	684	483	555	576	1.0										
LT3	527	338	383	089	095	009	619	428	458	641	675	1.0									
Cat	505	357	326	047	164	149	688	558	413	526	662	730	1.0								
Oti	744	322	525	244	234	-111	395	271	444	365	565	457	380	1.0							
Ari	912	382	610	255	284	-115	477	401	561	572	743	567	518	774	1.0						
MFW	592	215	482	155	175	-325	165	281	375	280	339	374	373	375	583	1.0					
Eng	853	338	498	263	324	-114	432	380	488	476	670	544	539	697	841	543	1.0				
Voc	816	308	547	214	302	-158	358	348	439	460	613	458	450	715	786	464	806	1.0			
L&N	657	424	420	248	174	006	647	480	467	528	676	651	593	524	691	379	723	556	1.0		
OI	758	357	367	290	208	-056	444	394	321	528	605	485	529	594	780	470	788	655	596	1.0	
EF	451	339	350	083	212	-007	697	475	417	532	585	614	614	374	506	298	444	424	504	473	1.0

* decimal points omitted in off-diagonals

INTERCORRELATIONS - FACTOR ANALYSIS VARIABLES - ESKIMOS 12½-15½ *

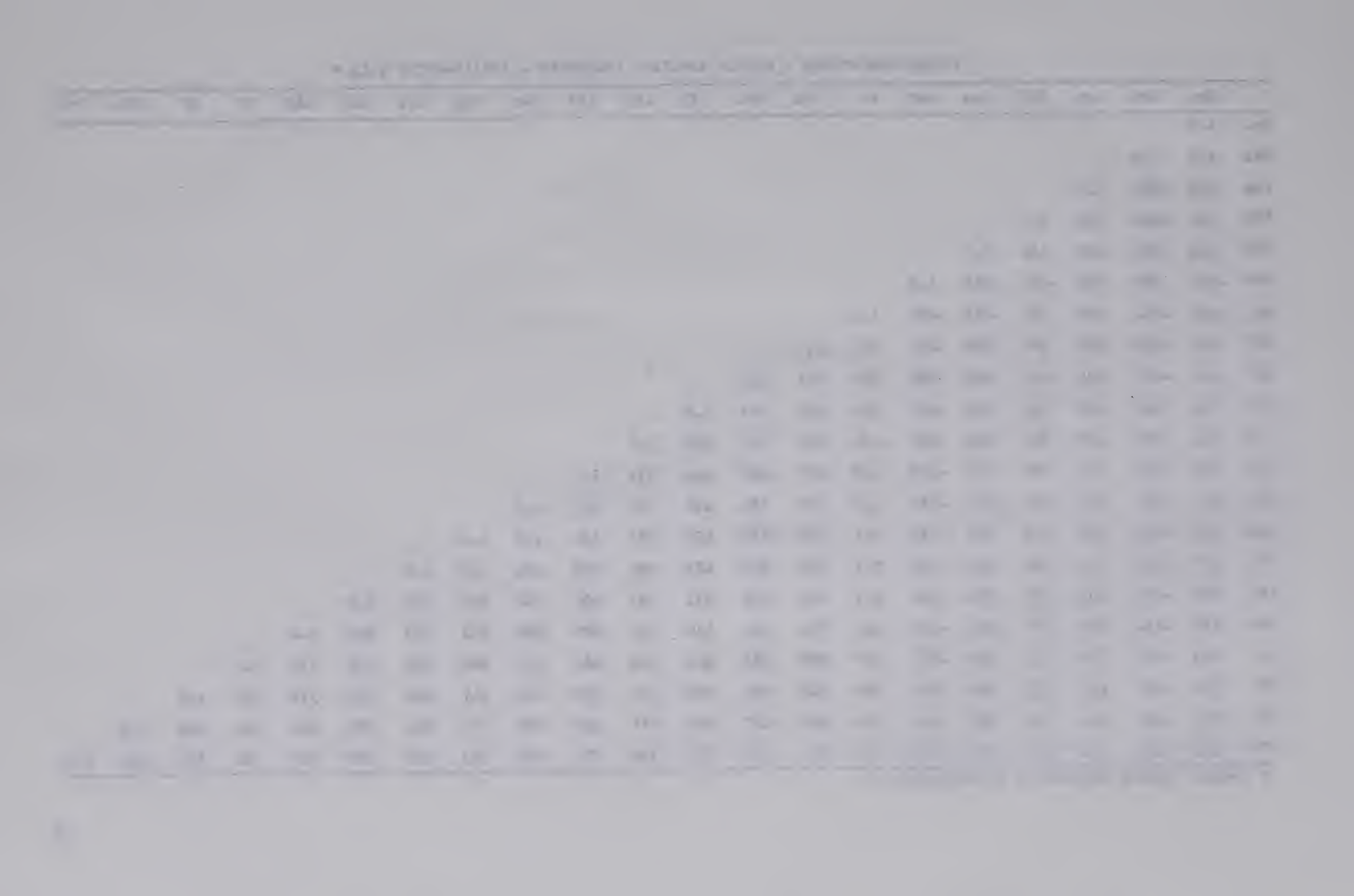
	Gr.	Age	TIS	Occ	Pl.	Sex	PM	SCR	MAC	LT1	LT2	LT3	Cat	Oti	Ari	MFW	Eng	Voc	L&N	OI	EF
Gr.	1.0																				
Age	515	1.0																			
TIS	840	475	1.0																		
Occ	056	087	090	1.0																	
Pl.	242	043	196	001	1.0																
Sex	-197	-036	-163	-047	040	1.0															
PM	644	250	522	022	264	-082	1.0														
SCR	292	033	139	159	159	-131	618	1.0													
MAC	336	185	181	-012	-038	-133	475	415	1.0												
LT1	483	229	371	-003	158	137	435	300	408	1.0											
LT2	650	184	467	073	326	005	601	463	339	528	1.0										
LT3	397	082	311	-011	271	054	545	431	397	490	658	1.0									
Cat	197	080	113	-073	190	134	522	426	301	376	448	555	1.0								
Oti	794	240	582	-069	297	-220	639	384	357	411	700	499	284	1.0							
Ari	871	367	711	035	274	-080	681	368	405	494	729	522	275	771	1.0						
MFW	332	111	223	092	188	-015	322	279	277	244	389	305	225	273	338	1.0					
Eng	831	269	681	-009	315	-219	574	312	341	445	658	498	197	809	799	280	1.0				
Voc	739	206	610	-001	321	-095	443	191	194	401	574	327	089	666	668	219	795	1.0			
L&N	660	283	496	-029	391	-091	670	404	345	495	711	594	483	720	741	313	736	483	1.0		
OI	588	212	551	-023	264	-095	431	216	233	298	530	439	268	628	638	178	739	596	620	1.0	
EF	544	206	366	005	227	049	674	519	414	414	597	566	546	612	592	245	565	521	608	518	1.0

* decimal points omitted in off-diagonals

INTERCORRELATIONS - FACTOR ANALYSIS VARIABLES - INDIAN-METIS 9-12 *

	Gr.	Sex	Age	TIS	Occ	Amt	PM	SCR	MAC	LT1	LT2	LT3	Cat	Rdg	Ari	Lan	L&N	OI	EF	CTL	CTN
Gr.	1.0																				
Sex	-119	1.0																			
Age	652	-034	1.0																		
TIS	904	-108	744	1.0																	
Occ	215	-065	-020	194	1.0																
Amt	-337	082	038	-302	-532	1.0															
PM	457	-014	305	350	-125	-207	1.0														
SCR	450	-190	071	307	186	-210	411	1.0													
MAC	499	-070	174	410	240	-281	382	441	1.0												
LT1	206	149	109	163	113	-226	555	170	415	1.0											
LT2	414	-106	146	260	232	-221	483	529	478	490	1.0										
LT3	302	-213	197	228	100	-276	608	297	470	554	531	1.0									
Cat	432	021	248	320	-005	-235	619	419	334	492	360	511	1.0								
Rdg	845	-046	468	742	215	-336	541	547	526	421	573	450	478	1.0							
Ari	870	-161	555	760	288	-364	545	516	560	423	604	478	484	910	1.0						
Lan	808	-239	503	735	196	-301	477	456	491	331	523	450	442	893	870	1.0					
L&N	520	-234	187	335	243	-306	544	566	429	430	769	688	466	711	712	660	1.0				
OI	783	-117	419	613	340	-374	467	606	565	348	648	465	475	818	858	758	737	1.0			
EF	589	-125	465	537	080	-181	649	525	273	363	525	338	423	647	631	565	552	549	1.0		
CTL	697	-126	354	500	228	-339	499	510	417	362	618	447	498	739	741	678	620	795	538	1.0	
CTN	542	025	371	403	157	-135	290	533	351	334	593	273	458	603	653	528	587	716	512	531	1.0

* decimal points omitted in off-diagonals



INTERCORRELATIONS - FACTOR ANALYSIS VARIABLES - INDIAN-METIS 12½-15½ *

	Gr.	Sex	Age	TIS	Occ	Amt	PM	SCR	MAC	LT1	LT2	LT3	Cat	Rdg	Ari	Lan	L&N	OI	EF	CTL	CTN
Gr.	1.0																				
Sex	047	1.0																			
Age	438	115	1.0																		
TIS	590	196	391	1.0																	
Occ	048	215	142	-021	1.0																
Amt	-152	-161	035	-085	011	1.0															
PM	531	166	119	175	001	-361	1.0														
SCR	444	-060	195	192	-005	-419	767	1.0													
MAC	148	-023	050	-031	-106	-258	441	470	1.0												
LT1	387	228	173	087	-026	-403	462	379	420	1.0											
LT2	503	072	221	211	-023	-286	583	703	301	399	1.0										
LT3	487	217	385	294	-058	-376	720	643	400	580	590	1.0									
Cat	428	359	306	196	081	-432	554	489	482	584	471	672	1.0								
Rdg	759	083	138	330	-128	-208	681	578	360	460	608	537	518	1.0							
Ari	602	386	123	306	012	-302	758	525	269	550	617	637	467	691	1.0						
Lan	671	-002	026	269	-155	-236	687	547	312	431	601	519	440	827	707	1.0					
L&N	598	-014	081	253	-112	-359	771	767	407	359	705	642	559	758	652	741	1.0				
OI	297	133	-039	066	093	-022	337	386	213	212	393	161	099	515	359	462	334	1.0			
EF	359	150	145	219	021	-257	596	591	483	316	543	545	635	415	422	324	574	210	1.0		
CTL	619	157	132	341	-033	-259	672	666	376	366	656	537	440	796	678	781	770	522	480	1.0	
CTN	366	310	193	130	-103	-238	554	426	229	336	257	524	548	507	365	447	465	229	410	486	1.0

* decimal points omitted in off-diagonals

APPENDIX C

MAC - OUTLINE OF DIRECTIONS

Total Time: Up to 60 minutes

The objective of MAC1, questions 1 to 10, is thorough understanding by each pupil of: (a) the principles involved in selecting the next two pieces to continue the pattern, and (b) the mechanics of selecting and recording answers. The pacing used by the teacher on these ten items should also help to standardize the pace at which the remaining items are done by the pupils. These first ten items are teaching items.

The test proper consists of MAC1, questions 11 to 20, and MAC2.

I. Distribution and Discussion of MAC1 Materials - (5 minutes)

1. Distribute Answer Sheets for MAC1. Write your first name and your last name, and the number 1 after your name.

WE WILL WRITE ONLY ON THIS ANSWER SHEET

2. Distribute MAC1 Booklets. Insert answer sheets as follows, teacher demonstrating at front of room:
   Booklets placed flat on desk, yellow side up, staples at right.
   Open the yellow cover wide to the right, and place answer sheet on top of yellow cover, with lines upward on the right.
   Check that all have answer sheets correctly inserted.
   Close booklets, white side up, and staples to left.

3. Distribute Choice Sheets. Place on desk just above booklet. Brief general explanation of choice sheets, bringing out by questioning: (a) Rows of pieces, (b) Some Blue, Red, Green, (c) Some Circles, Squares, Triangles, (d) Some Large, Medium, Small (Three bears reference for younger grades). Use the above words, while pointing to examples on the choice sheet. Each piece has a name. What is the name of this piece? (B1). What is the name of this piece? (B7). What is the name of this piece?

II. Teaching MAC1, Questions 1 to 10 (15 minutes)

4. Work Question 1 with class.
   Now, open your booklet to the first page. In the top row of your book here, see the row of blue large circles. (Demonstrate). They make a pattern going across the page this way. (Demonstrate, sweeping finger across page, left to right).

Your job is to add TWO more pieces to the end of the pattern where the blanks are on your answer sheet here (point), to make the pattern keep on in the right way.

You may choose the pieces from the ones on this choice sheet. When you know which pieces are needed to keep on the pattern in the right way here in your booklet, write their numbers on the answer sheets in the right blanks. Any answer piece may be used more than once.

For Question 1, what piece number is needed to keep the pattern going on? Put your finger on the piece on this choice sheet that is needed to keep the pattern going on. Its number is B1, so write B1 on your answer sheet here.

Then B1 is needed again, so write it again on your answer sheet here. WE WRITE ONLY ON THE ANSWER SHEET.

Check that all have completed Question 1 correctly, assisting where necessary.

5. Work Question 2 with class.
Now look at Question 2. The TWO pices needed to keep the pattern going are a red large triangle, and another red large triangle, so what numbers shall we write on our answer sheet? R7, R7. Write in R7, R7. Check that all have completed Question 2 correctly, assisting where necessary.

6. Work Question 3 with class.
Now Question 3. The next piece for question 3 will be a large circle. What color should it be? Blue. So what number shall we put in the next blank? B1. Write in B1.

And then what color should the next large circle in this row be? Red. So what number shall we write in next? R1. Write in R1. Check, and help any pupils in difficulty. Several may have difficulty with Question 3.

7. Work Questions 4 and 5 with class, similarly.

8. Now have pupils turn the page, and proceed right on by themselves with Question 6 to 10. This is still a teaching page.
Check carefully and help any individuals having difficulty with question 6 to 10, noting these points:
   (a) Each pupil should understand the principles involved in selecting the next TWO pieces to continue the pattern.

(b) Each pupil should understand the mechanics of selecting and recording answers. Watch for reversals in recording.
(c) Attempt to establish appropriate pacing, so that pupils are neither dashing through cavalierly, nor dawdling.

III. Testing MAC1, Questions 11 to 20 (15 minutes)

9. Now turn over to the next page, and go on starting with Question 11, right to the end of the booklet, on your own. Check for understanding of mechanics of turning pages and recording, and encourage dawdlers, but give no further help with content of test, except to say if necessary, "Can you find the numbers of the next two pieces needed to keep the pattern going on in the right way? Then write their numbers here."

Allow up to 15 minutes working time for Questions 11 to 20.

IV. Collection of MAC1 Materials, and Distribution of MAC2 Materials (3 minutes)

10. Remove Answer Sheets from booklets. Write your first name and your last name on the answer sheet. Collect answer sheets.

11. Collect MAC1 Booklets. Leave choice sheets on desks.

12. A slight break (standing and stretching at desks) may be useful here.

13. Distribute Answer Sheets for MAC2. (Same kind of sheets as MAC1). Write your first name and your last name, and the number 2 after it.

14. Distribute MAC2 Booklets. Insert answer sheets as before, teacher demonstrating and checking.

V. Testing MAC2 (20 minutes)

12. Have pupils work through all of MAC2 at their own pace, watching these points:
(a) It may be necessary to give further help with Questions 1 and 2 to get pupils started. The color in Question 3 is intended to be Blue.
(b) Thereafter give no help except to say if necessary, "Can you find the numbers of the next two pieces needed to keep the pattern going on in the right way? Then write their numbers here."
(c) This is a power test. Allow up to 20 minutes working time for MAC2.

(d) Watch any making no progress or dawdling and urge quietly.

VI. Collection of MAC2 Materials (2 minutes)

16. Remove Answer Sheets from Booklets. Write your first name and your last name on the answer sheet. Collect answer sheets.

17. Collect MAC2 Booklets

18. Collect Choice Sheets. (Except for Grades 7 and 8).

19. Compliment Class.

VII. Grades 7 and 8 Only -- Testing MAC3

20. Distribute MAC3 Answer Sheets. Write names. Distribute MAC3 Booklets. Insert Answer Sheets.

21. Proceed directly to Testing MAC3. Allow up to 20 minutes working time.

22. Collect Answer Sheets. Collect Booklets, Collect Choice Sheets. Compliment class.

APPENDIX D

MAC ITEMS - DIFFICULTY LEVELS AND POINT

BISERIAL r'S WITH TOTAL TEST SCORES

Item No.	Esk 9-12		Esk 12½-15½		I-M 9-12		I-M 12½-15½		W 9-12		W 12½-15½	
	r	d	r	d	r	d	r	d	r	d	r	d
1	.41	.91	.15	.96	.66	.89	.34	.96	.18	.96	.06	.98
2	.52	.84	.50	.95	.55	.64	.44	.87	.60	.92	.33	.98
3	.61	.79	.50	.91	.52	.84	.01	.98	.51	.94	.75	.96
4	.68	.83	.64	.94	.67	.77	.47	.96	.64	.95	.62	.94
5	.58	.47	.63	.86	.71	.41	.54	.85	.63	.80	.82	.94
6	.58	.77	.44	.91	.65	.66	.53	.91	.40	.91	.41	.87
7	.64	.79	.48	.89	.61	.66	.06	.96	.49	.87	.63	.91
8	.67	.56	.67	.83	.71	.38	.64	.80	.73	.79	.70	.85
9	.56	.82	.48	.93	.66	.73	.25	.94	.66	.96	.72	.94
10	.66	.40	.59	.70	.62	.39	.47	.85	.72	.70	.60	.87
11	.18	.39	.28	.54	.18	.39	.20	.35	.30	.38	.50	.66
12	.71	.59	.58	.71	.63	.36	.38	.74	.70	.77	.73	.66
13	.56	.41	.55	.80	.62	.38	.52	.54	.60	.61	.54	.72
14	.50	.09	.55	.20	.00	.00	.49	.13	.27	.06	.26	.19
15	.34	.11	.39	.18	.24	.02	.28	.07	.19	.01	.37	.09
16	.27	.08	.49	.15	.38	.09	.45	.15	.21	.04	.44	.19
17	.19	.05	.29	.09	.00	.00	.27	.09	.19	.01	.35	.04
18	.00	.00	.31	.05	.00	.00	.00	.00	.16	.01	.16	.02

APPENDIX E

MEANS AND STANDARD DEVIATIONS OF VARIABLES

ON WHICH NO TRANSFORMATIONS WERE MADE

ESKIMOS 9-12

Variable	Mean	S.D.
Grade in School	3.34	1.11
Age in months	129.53	11.56
Total Time in School in years	4.12	1.30
Occupation of Parent (Blishen Index Score)	37.91	6.11
Plans for Age 20 (first digit only of Blishen Index score for occupation chosen)	4.01	1.18
Sex (1 for girls, 2 for boys)	1.70	0.46

ESKIMOS 12½-15½

Variable	Mean	S.D.
Grade	5.40	1.58
Age	163.95	11.31
Time in School	5.87	1.86
Occupation of Parent	38.24	6.76
Plans for Age 20	4.11	1.22
Sex	1.45	0.50

Note: All other variables normalized and transformed to a mean of 50 and standard deviation of 10.

MEANS AND STANDARD DEVIATIONS OF VARIABLES

ON WHICH NO TRANSFORMATIONS WERE MADE

INDIAN-METIS 9-12

Variable	Mean	S.D.
Grade in School	3.93	1.15
Age in months	127.39	11.66
Total Time in School in years	4.16	1.94
Occupation of Parent (Blishen Index Score)	38.64	7.19
Amount of Language other than English used in the home[a]	2.43	0.75
Sex (1 for girls, 2 for boys)	1.55	0.50

INDIAN-METIS 12½-15½

Variable	Mean	S.D.
Grade	6.52	1.17
Age	166.24	11.08
Time in School	7.17	1.47
Occupation of Parent	37.94	6.12
Amount of Language other than English	2.30	0.76
Sex	1.52	0.50

Note: All other variables normalized and transformed to a mean of 50 and standard deviation of 10.

[a]Coded 1 for English only, 2 for small part of the time, 3 for most of the time and 4 for all of the time.

APPENDIX F

EXAMPLE ITEMS SIMILAR TO THOSE OF THE MAC2